Journal of Engineering and Applied Sciences 12 (8): 1992-1998, 2017
ISSN: 1816-949X

In Silico Potential Vaccine Against Outer Membrane Protein 2 in *Haemophilus parasuis*

Nazgol Emamian, Fateme Sefid, Vahide Saeidjavan, Sepide Akhgari and Hamed Karamizade
Departmant of Biology, Science and Art University, Yazd, Iran

Abstract: *Haemophilus parasuis* is a gram-negative bacterium belonging to the *Pasteurellaceae* family. This organism is an important respiratory-tract pathogen in swine and the etiological agent of porcine polyserositis, meningitis and arthritis syndrome known as Glasser's disease. Among the 15 known serovars, serotype 5 shows high virulence and is one of the most prevalent serotypes. Attempts to control the *H. parasuis* infection are hindered by a lack of thorough knowledge of the virulence factors and protective antigens of the bacterium, the existence of diverse genetic make-ups and the evolution of multidrug-resistant strains. Initial studies about the immune response developed against *H. parasuis* have detected antibodies to Outer Membrane Proteins (OMPs) but not against lipopolysaccharide or capsule, suggesting that the OMPs are more immunogenic than other components of bacteria. In recent years, interest has shifted toward protein based vaccines. It has been shown that recombinant vaccines based on OMPs provided partial protection against challenge with *H. parasuis*.

Key words: *Haemophilus parasuis*, OMP, bioinformatic, immune response, partial protection

INTRODUCTION

Haemophilus parasuis is a gram negative bacterium belonging to the Pasteurellaceae family. This organism is an important respiratory-tract pathogen in swine and the etiological agent of porcine polyserositis, meningitis and arthritis syndrome known as Glasser's disease (Bigas *et al*., 2010). With the recent changes in production methods, diseases caused by *H. parasuis* have become increasingly significant world wide (Hill *et al*., 2003).

Among the 15 known serovars, serotype 5 shows high virulence and is one of the most prevalent serotypes. Attempts to control the *H. parasuis* infection are hindered by a lack of thorough knowledge of the virulence factors and protective antigens of the bacterium the existence of diverse genetic make-ups and the evolution of multidrug resistant strains (Oliveira and Pijoan, 2004). Initial studies about the immune response developed against *H. parasuis* have detected antibodies to Outer Membrane Proteins (OMPs) but not against lipopolysaccharide or capsule, suggesting that the OMPs are more immunogenic than other components of bacteria (Miniats *et al*., 1991). In recent years, interest has shifted toward protein based vaccines. It has been shown that recombinant vaccines based on OMPs provided partial protection against challenge with *H. parasuis* (Martin *et al*., 2009).

Therefore, it is necessary to identify novel and more efficient immunoprotective antigens contributing to the development of a monovalent or a multivalent subunit vaccine that can protect pigs against *H. parasuis* infection. Identifying immunogens is crucial for vaccine development and recent progresses in genomic and proteomic technologies have made it possible to perform global profiling of immunogenic proteins for bacterial pathogens (Chen *et al*., 2004; Krah and Jungblut, 2004). The OMPs of Gram-negative bacteria have been considered important targets for vaccine development and diagnostic candidates (Nally *et al*., 2005).

The property of an antigen to bind specifically complementary antibodies is known as the antigen's antigenicity; likewise, the ability of an antigen to induce an immune response is called its immunogenicity. Attempts should be made to discover peptides that could mimic protein epitopes and possess the same immunogenicity as the whole protein. New genome analysis tools based on bioinformatics and immunoinformatics approaches help us select suitable antigens or epitopes directly from the genomes of pathogens in order to design a vaccine. These tools could be employed for epitope selection and vaccine design. Moreover, prediction of protein structures is one of their wide applications (Kofeiti *et al*., 2016; Boyce *et al*., 2006; Bastani and Sefid, 2016; Darbandian and Sefid, 2017; Sefid *et al*., 2013, 2015, 2016).

Subsequently, theoretical methods for epitope prediction have been developed leading to synthesis of

Corresponding Author: Fateme Sefid, Departmant of Biology, Science and Art University, Yazd, Iran

such peptides that are important for development of immunodiagnostic tests and vaccines. The present study was designed to in silico resolving the major obstacles in the control or in prevention of the diseases caused *Haemophilus parasuis*. We exploited bioinformatic tools to better understanding and characterizing the OMP2 structure and select appropriate regions as effective B cell epitops.

MATERIALS AND METHODS

Sequence availability: The OMP2 protein sequence with accession No. AGO16427.1 and GI: 513128207 acquired from NCBI at http://www.ncbi.nlm.nih.gov/protein were saved in FASTA format for further analyses.

Homology search: The sequences served as a query for protein BLAST at http://blast.ncbi.nlm.nih.gov/Blast.cgi against non redundant protein database. Probable putative conserved domains of the query protein were also searched for at the above address.

Primary sequence analysis: Protparam online software at http://expasy.org/tools/protparam.html was employed for estimation and determination of properties such as molecular weight, theoretical pI, amino acid composition, total number of negatively and positively charged residues, instability index and aliphatic index.

Subcellular localization: Subcellular localization of the protein was predicted by CELLO at http://cell.life.nctu.edu.tw/

3D structure prediction: The Swiss-model workspace at http://swissmodel.expasy.org/is a web-based integrated service dedicated to protein structure homology modelling. It assists and guides the user in building protein homology models at different levels of complexity.

Building a homology model comprises four main steps: identification of structural template(s), alignment of target sequence and template structure(s), model building and model quality evaluation. These steps can be repeated until a satisfying modelling result is achieved. Each of the four steps requires specialized software and access to up-to-date protein sequence and structure databases.

Models evaluations: All 3D models of the proteins built, were qualititatively estimated by GMQE and QMEAN4 scores.

Single-scale amino acid properties assay: IEDB at http://tools.immuneepitope.org/tools/bcell/iedb-input parameters such as hydrophilicity, flexibility, accessibility, turns and antigenic propensity of polypeptide have been correlated with the location of B cell epitopes. This has led to a search for empirical rules that would allow the position of B cell epitopes to be predicted from certain features of the protein sequence.

B cell epitope prediction: Ellipro at http://tools.immuneepitope.org/tools/Ellipro/tutorial.jsp predicts linear and discontinuous antibody epitopes.

RESULTS AND DISCUSSION

Sequence availability and homology search: The protein sequence with 344 residue obtained from NCBI and saved in FASTA format. Protein sequence serving as query for BLAST produced a set of sequences as the highest similar sequence.

BLAST search revealed numerous hits to the OMP2 subunit sequence (Fig. 1). Most of hits were of Haemophilus. Putative conserved domain were detected within this sequence. Most of the sequences belong to OM-Channels super family. These outer membrane channels share a beta-barrel structure that differ in strand and shear number. Classical (gram-negative) porins are non-specific channels for small hydrophillic molecules and form 16 beta-stranded barrels which associate as trimers. Maltoporin-like channels have specificities for various sugars and form 18 beta-stranded barrels which associate as trimers. Ligand-gated protein channels cooperate with a TonB associated inner membrane complex to actively transport ligands via the proton motive force and they form monomeric, barrels. The 150-200 N-terminal residues form a plug that blocks the channel from the periplasmic end.

Primary sequence analysis: The protein sequence served as input for the computation of various physical and chemical parameters. The computed parameters included the molecular weight, theoretical pI, instability index, aliphatic index and grand average of hydropathicity (indicates the solubility of the proteins: positive GRAVY (hydrophobic), negative GRAVY (hydrophilic) are summarized in Table 1 and 2.

Extinction coefficients: Extinction coefficients are in units of M^{-1} cm^{-1} at 280 nm measured in water:

- Ext. coefficient 42290
- Abs 0.1% (= 1 g/L) 1.099

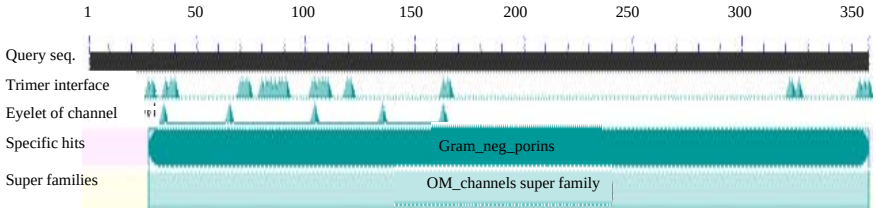


Fig. 1: Putative conserved domains have been detected

Table 1: Amino acid composition

Protein	Percentage
Ala (A) 32	08.9
Arg (R) 10	02.8
Asn (N) 24	06.7
Asp (D) 22	06.1
Cys (C) 0	00.0
Gln (Q) 9	02.5
Glu (E) 15	04.2
Gly (G) 47	13.1
His (H) 7	02.0
Ile (I) 13	03.6
Leu (L) 21	05.9
Lys (K) 37	10.3
Met (M) 2	00.6
Phe (F) 14	03.9
Pro (P) 2	00.6
Ser (S) 21	05.9
Thr (T) 27	07.5
Trp (W) 2	00.6
Tyr (Y) 21	05.9
Val (V) 32	08.9
Pyl (O) 0	00.0
Sec (U) 0	00.0

No. of amino acids: 358; molecular weight: 38492.9; theoretical pI: 9.26

Table 2: Atomic composition

Atoms	Composition
Carbon (C)	1710
Hydrogen (H)	2670
Nitrogen (N)	474
Oxygen (O)	535
Sulfur (S)	2

Formula: $C_{1710}H_{2670}N_{474}O_{535}S_2$; Total number of atoms: 5391; total number of negatively charged residues (Asp+Glu): 37; total number of positively charged residues (Arg+Lys): 47

Estimated half-life: The N-terminal of the sequence considered is M (Met). The estimated half-life is 30 h (mammalian reticulocytes *in vitro*):

- >20 h (yeast *in vivo*)
- >10 h (*Escherichia coli in vivo*)

Instability index: The Instability Index (II) is computed to be 15.51. This classifies the protein as stable:

- Aliphatic index: 71.90
- Grand average of hydropathicity (GRAVY) -0.475

Subcellular localization: OMP2 subcellular localization predicted by CELLO was outer membrane with the highest reliability index (4.106).

CELLO prediction:

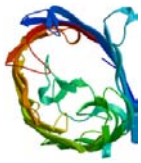
- Outer membrane = 4.106*
- Extracellular = 0.568
- Periplasmic = 0.269
- Cytoplasmic = 0.037
- Inner membrane = 0.020

3D structure prediction: Building a homology model comprises four main steps: identification of structural template(s), alignment of target sequence and template structure(s), model building and model quality evaluation. These steps can be repeated until a satisfying modelling result is achieved. Each of the four steps requires specialized software and access to up-to-date protein sequence and structure databases. Swiss model software recruited for homology modeling introduced 3 model. All the models were selected for further analyses.

Models evaluations: The 3D models estimated qualititatively by tow servers revealed that there was a consensus on a single model. Results are shown in Table 1. QMEAN is a composite scoring function for the estimation of the global and local model quality. QMEAN consisting of four structural descriptors: the local geometry is analysed by a torsion angle potential over three consecutive amino acids. Two pairwise distance dependent potentials are used to assess all atom and C-beta interactions. A solvation potential describes the burial status of the residues. The pseudo energies returned from the four structural descriptors and the final QMEAN4 score get directly related to what we would expect from high resolution X-ray structures of similar size using a z-score scheme. The score of a model in also shown in relation to a set of high-resolution PDB structures (z-score).

The plot relates the obtained global QMEAN 4 value to scores calculated from a set of high-resolution X-ray structures. Local estimates of the model quality based on the QMEAN scoring function are shown as per-reside plot. Each residue is assigned a reliability score between 0 and 1, describing the expected similarity to the native structure. Higher numbers indicate higher reliability of the residues. GMQE (Global Model Quality Estimation) is a quality estimation which combines properties from the target-template alignment. The resulting GMQE score is

Table 3: OMP2 predicted 3D models estimated qualititatively

Models	Figures	Template	Seq. identity (%)	Seq. similarity	Coverage	GMQE	QMEAN4
1		3nsg.1.A	17.80	0.30	0.86	0.52	-11.72
2		3nsg.1.C	17.80	0.30	0.86	0.52	-12.16
3		3nsg.1.B	17.80	0.30	0.86	0.52	-12.35

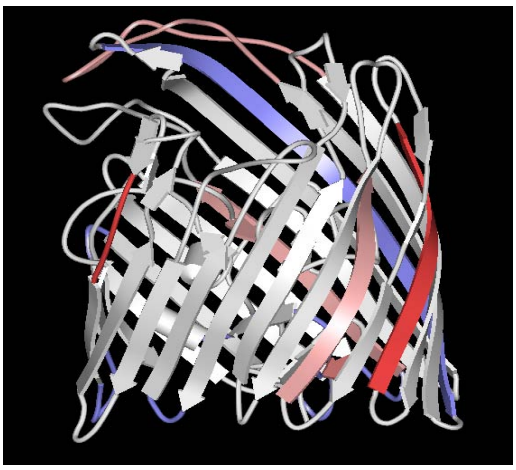


Fig. 2: PalA 3D structure model

expressed as a number between zero and one, reflecting the expected accuracy of a model built with that alignment and template. Higher numbers indicate higher reliability. Once a model is built, the GMQE gets updated for this specific case by also taking into account the QMEAN4 score of the obtained model in order to increase reliability of the quality estimation. The best 3D model is shown in Fig. 2 and Table 3.

Single-scale amino acid properties assay: IEDB server predicts several properties such as hydrophilicity, accessibility, antigenicity, flexibility and beta turn secondary structure in the protein sequence. Propensity scale methods assign a propensity value to each amino acid which measures the tendency of an amino acid to be part of a B cell epitope (as compared to the background). To reduce fluctuations the score for each target amino acid residue in a query sequence is computed as the average of the propensity values of the amino acids in a sliding window centered at the target residue. hydrophilicity, accessibility, antigenicity, flexibility and secondary structure properties have fundamental role in B cell epitope prediction. Relying on just one of these properties, reliable results could not be achieved. Results are shown in Fig. 3.

Parameters such as hydrophilicity, flexibility, accessibility, turns, exposed surface, polarity and antigenic propensity of polypeptides chains have been correlated with the location of continuous epitopes. This has led to a search for empirical rules that would allow the position of continuous epitopes to be predicted from certain features of the protein sequence. All prediction calculations are based on propensity scales for each of the 20 amino acids. Each scale consists of 20 values assigned to each of the amino acid residues on the basis of their relative propensity to possess the property described by the scale.

When computing the score for a given residue i, the amino acids in an interval of the chosen length, centered around residue i are considered. In other words, for a window size n the i-(n-1)/2 neighboring residues on each side of residue i were used to compute the score for residue i unless specified the score for residue i is the average of the scale values for these amino acids. In general, a window size of 5-7 is appropriate for finding

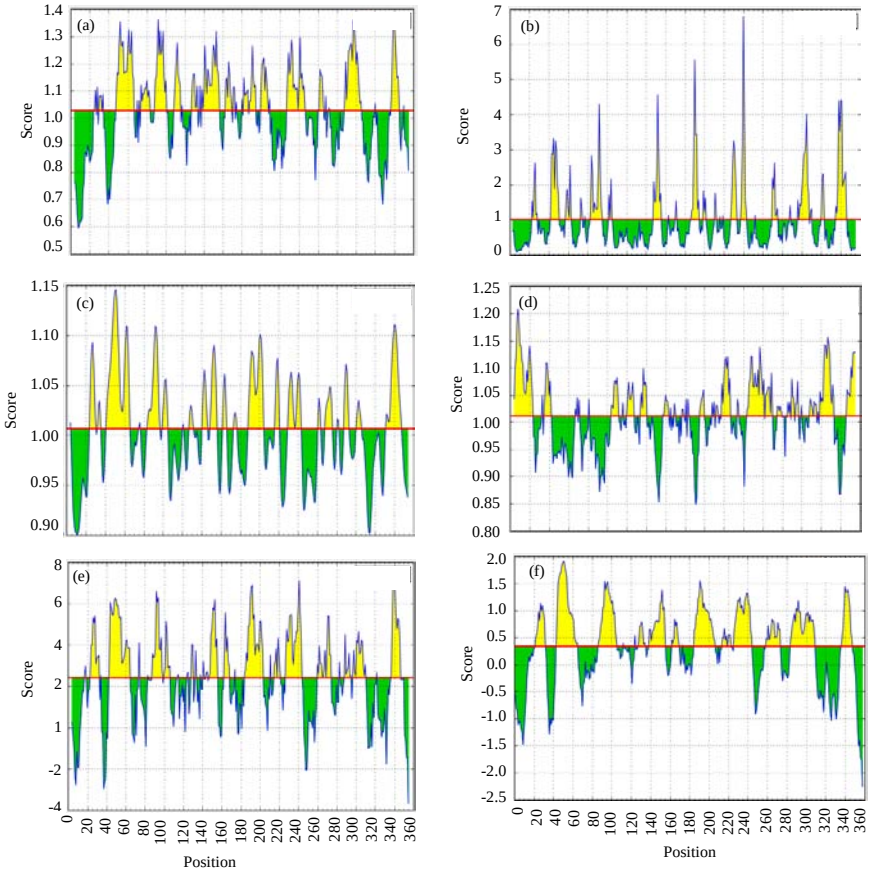


Fig. 3: IEDB linear B cell epitope prediction results for OMP2 protein: a) Hydrophilicity; b) Accessibility; c) Antigenicity; d) Flexibility and e) Beta turn and f) Bepipred linear epitope

Table 4: Linear eitopes pedicted by Ellipro

Start	End	Peptide	No. of residues	Score
20	32	VTVYENEGTKVDF	13	0.842
68	80	ISIKHNINENLYG	13	0.808
285	293	FDVTPKSGV	9	0.780
211	221	YTAKIAESQSV	11	0.757
229	239	HDDYKSGAVNK	11	0.746
301	309	TYKDKAYKA	9	0.742
88	100	LGRNSKNDAGWGD	13	0.736
186	196	VANERDNKGEV	11	0.736
36	62	LRLLLEEQATKEKGQSSTGGHTNLKNN	27	0.684
320	328	YKLHKQVVT	9	0.684
334	342	LIKNKDSNN	9	0.655
108	115	VGLGGYGH	8	0.653

regions that may potentially be antigenic. On Fig. 4, the Y-axis depicts for each residue the correspondent score (averaged in the specified window) be it a BepiPred score or a residue score on the Karplus and Schulz flexibility scale while the X-axes depicts the residue positions in the sequence. Table 4 provide values of calculated scores for each residue. The larger score for the residues might be interpreted as that the residue might have a higher probability to be part of epitope (those residues are colored in yellow on the graphs).

Prediction of B cell epitopes by integrated strategy: Four linear along with 11 discontinuous B cell epitopes were predicted by Ellipro Software (Table 4 and 5). Two discontinuous and 2 linear epitopes with the highest PI (Protrusion Index) are shown in Fig. 4.

Table 5: Dscontinuous eitopes pedicted by Ellipro

Residues	No. of residues	Scores
A:E24, A:N25, A:E26	3	0.972
A:V20, A:T21, A:V22, A:Y23, A:G27, A:T28, A:K29, A:V30, A:D31, A:F32, A:L36, A:L38, A:L40, A:E42, A:Q43, A:A44, A:T45, A:K46, A:E47, A:K48, A:G49, A:Q50, A:S51, A:S52, A:T53, A:G54, A:G55, A:H56, A:T57, A:N58, A:L59, A:K60, A:N61, A:N62, A:I70, A:K71, A:H72, A:N73, A:I74, A:N75, A:E76, A:N77, A:L78, A:Y79, A:G80, A:L88, A:G89, A:R90, A:N91, A:S92, A:K93, A:N94, A:D95, A:A96, A:G97, A:W98, A:G99, A:D100, A:V108, A:G109, A:L110, A:G111, A:G112, A:Y113, A:G114, A:H115, A:I117, A:Y172, A:T173, A:G174, A:I175, A:E176, A:G177, A:L178, A:Y211, A:T212, A:A213, A:K214, A:I215, A:A216, A:E217, A:S218, A:Q219, A:S220, A:V221, A:L249, A:Y251, A:V252, A:N253, A:A254, A:K273, A:F285, A:D286, A:V287, A:T288, A:P289, A:K290, A:S291, A:G292, A:V293, A:T301, A:Y302, A:K305, A:A306, A:Y307, A:K308, A:A309, A:H312, A:Y320, A:K321, A:L322, A:H323, A:K324, A:Q325, A:V326, A:T328, A:L334, A:I335, A:K336, A:N337, A:K338, A:D339, A:S340, A:N341, A:N342, A:V345, A:T346, A:D347, A:Q348, A:A349, A:L350, A:G351, A:V352, A:G353, A:L354, A:R355, A:V356, A:L357	138	0.713
A:K155, A:G156, A:F157, A:D158, A:V186, A:A187, A:N188, A:E189, A:R190, A:D191, A:N192, A:K193, A:G194, A:E195, A:V196, A:S200, A:T201, A:K202, A:S203, A:D231, A:Y232, A:K233, A:S234, A:G235, A:A236, A:V237, A:N238, A:K239, A:V267, A:K268, A:T269, A:E274, A:K275, A:K303, A:D304	35	0.672

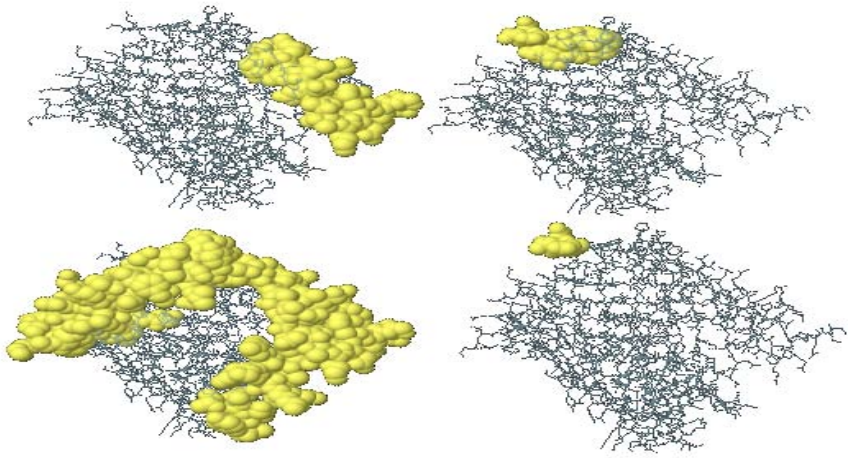


Fig. 4: The 2 linear (up) and 2 discontinuous (down) epitopes with the highest PI score predicted by Ellipro server are shown. Epitopes mapped on 3D models using discovery studio Visualizer 2.5.5 Software

CONCLUSION

The present study was designed to in silico resolving the major obstacles in the control or in prevention of the diseases caused *Haemophilus parasuis*. We exploited bioinformatic tools to better understanding and characterizing the OMP2 structure and select appropriate regions as effective B cell epitops.

REFERENCES

Bastani, A. and F. Sefid, 2016. Evaluation of Leptospira interrogans outer membrane protein Ompl37 linear and conformational epitopes. Intl. J. Biol. Biotechnol., 7: 586-593.

Bigas, A., M.E. Garrido, I. Badiola, J. Barbe and M. Llagostera, 2010. Colonization capacity and serum bactericidal activity of Haemophilus parasuis thy mutants. Intl. Microbiol., 9: 297-301.

Boyce, J.D., P.A. Cullen, V. Nguyen, I. Wilkie and B. Adler, 2006. Analysis of the Pasteurella multocida outer membrane sub proteome and its response to the in vivo environment of the natural host. Proteomics, 6: 870-880.

Chen, Z., B. Peng, S. Wang and X. Peng, 2004. Rapid screening of highly efficient vaccine candidates by immunoproteomics. Proteomics, 4: 3203-3213.

Darbandian, P. and F. Sefid, 2017. Evaluation of OMP A (Outer Membrane Protein A) linear and conformational epitopes in Acinetobacter baumannii. Intl. J. Adv. Biotechnol. Res., 7: 994-1003.

Hill, C.E., D.S. Metcalf and J.I. MacInnes, 2003. A search for virulence genes of Haemophilus parasuis using differential display RT-PCR. Vet. Microbiol., 96: 189-202.

Kofeiti, A., M. Kalantar and F. Sefid, 2016. In silico determination and validation of Oprl structure and B cell epitopes as vaccine candidate in Pseudomonas aeruginosa. Intl. J. Pharm. Technol., 8: 14314-14327.

Krah, A. and P.R. Jungblut, 2004. Immunoproteomics. In: Molecular Diagnosis of Infectious Diseases, Decler, J. and R. Udo (Eds.). Humana Press, New York, USA., ISBN:978-1-58829-221-6, pp: 19-32.

Martin, D.L.F.A.J., C.B.G. Martin, C.P. Martinez, M.J.G. Iglesias and F. Tejerina et al., 2009. Effect of different vaccine formulations on the development of glassers disease induced in pigs by experimental Haemophilus parasuis infection. J. Comp. Pathol, 140: 169-176.

Miniats, O.P., N.L. Smart and S. Rosendal, 1991. Cross protection among Haemophilus parasuis strains in immunized gnotobiotic pigs. Can. J. Vet. Res., 55: 37-41.

Nally, J.E., J.P. Whitelegge, R. Aguilera, M.M. Pereira, D.R. Blanco and M.A. Lovett, 2005. Purification and proteomic analysis of outer membrane vesicles from a clinical isolate of Leptospira interrogans serovar Copenhageni. Proteomics, 5: 144-152.

Oliveira, S. and C. Pijoan, 2004. Haemophilus parasuis: New trends on diagnosis, epidemiology and control. Vet. Microbiol., 99: 1-12.

Sefid, F., I. Rasooli and A. Jahangiri, 2013. In silico determination and validation of Acinetobacter baumannii utilization a structure and ligand binding site. Biomed Res. Intl., 2013: 1-14.

Sefid, F., I. Rasooli and Z. Payandeh, 2016. Homology modeling of a Camelid antibody fragment against a conserved region of Acinetobacter baumannii biofilm associated protein (Bap). J. Theor. Biol., 397: 43-51.

Sefid, F., I. Rasooli, A. Jahangiri and H. Bazmara, 2015. Functional exposed amino acids of BauA as potential ImmunoGen against Acinetobacter baumannii. Acta Biotheor., 63: 129-149.